Published in final edited form as:
Nat Biotechnol. 2013 April ; 31(4): 331–334. doi:10.1038/nbt.2539.

Protein interaction discovery using parallel analysis of translated ORFs (PLATO)

Jian Zhu[#1,4,9], **H. Benjamin Larman**[#1,2,3,4,9], **Geng Gao**[1,4,9], **Romel Somwar**[7], **Zijuan Zhang**[8], **Uri Laserson**[2,4,5], **Alberto Ciccia**[1,4,9], **Natalya Pavlova**[1,4,9], **George Church**[2,4], **Wei Zhang**[8], **Santosh Kesari**[6], and **Stephen J. Elledge**[1,4,9,*]

[1]Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA

[2]Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA

[3]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA

[4]Department of Genetics, Harvard University Medical School, Boston, MA

[5]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA

[6]Division of Neuro-Oncology, Translational Neuro-Oncology Laboratories, Department of Neurosciences, U.C. San Diego, Moores Cancer Center, La Jolla, CA

[7]Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY

[8]Department of Chemistry, University of Massachusetts Boston, Boston, MA

[9]Howard Hughes Medical Institute

[#] These authors contributed equally to this work.

Abstract

Identifying physical interactions between proteins and other molecules is a critical aspect of biological analysis. Here we describe PLATO, an *in vitro* method for mapping such interactions by affinity enrichment of a library of full-length open reading frames displayed on ribosomes, followed by massively parallel analysis using DNA sequencing. We demonstrate the broad utility of the method for human proteins by identifying known and previously unidentified interacting partners of LYN kinase, patient autoantibodies, and the small-molecules gefitinib and dasatinib.

Several methods have been developed to characterize the specificities of protein-binding molecules. Fragmented cDNA libraries can be displayed on phage, bacteria or yeast, such

[*]Correspondence to: selledge@genetics.med.harvard.edu.

Contributions S.J.E. and H.B.L. conceived and supervised the project. pRD human ORFeome library was constructed by J.Z., and characterized by J.Z. and H.B.L. The PLATO protocol was developed by H.B.L. and J.Z. Clinical evaluations and patient sample acquisitions were performed by S.K. Statistical analysis was performed by U.L. under the supervision of G.M.C. R.S. provided gefitinibconjugated beads. PLATO candidates were confirmed by J.Z. and G.G. A.C. provided support for the validation of LYN binding candidates. N.P. provided support for the validation of PND autoantigen candidates. Z.Z. and W.Z. provided biotin-dasatinib. The manuscript was prepared by H.B.L. and J.Z., and edited by S.J.E.

that bait-binding proteins are identified by affinity enrichment or fluorescence-activated cell sorting. These approaches have limited success displaying long polypeptides and suffer from highly skewed clonal abundances, only the minority of which express coding sequences in the correct reading frame.[1] A variety of two-hybrid and split-reporter techniques have been developed to assay the binding of full length open reading frames (ORFs) within bacterial or eukaryotic cellular cytoplasm,[2] but these systems are limited to analyses of bait molecules that can be presented within the cell, and thus are not suitable for antibody analysis or drug target identification. More recently, protein microarrays have been used for these purposes,[3] but arraying typically requires individual proteins be purified and immobilized, resulting in significant costs and various degrees of protein denaturation.

To address these limitations, we developed a method called PLATO (ParalleL Analysis of Translated ORFs) that combines *in vitro* display of full length, solution phase proteins with cost effective analysis by high throughput DNA sequencing. We demonstrate the utility of PLATO by performing diverse interaction screens against the human ORFeome, a normalized collection of 15,483 cloned cDNAs housed within the Gateway recombination cloning system.[4]

PLATO employs ribosome display to express an ORF library. Ribosome display is a technique used to prepare a library of mRNA molecules that remain tethered to the proteins they encode via noncovalent interactions with a ribosome, due to absence of stop codons required for polypeptide release.[5] In contrast to alternative display platforms, ribosome display imposes minimal constraints upon the length or composition of proteins that can be efficiently displayed.

We constructed a Gateway cloning-compatible ribosome display “destination” vector (pRD-DEST; Supplementary Figure 1), which can be used as a recipient for a normalized pool of ORF-containing “entry” clones. After recombination, PCR amplification yields a linear template lacking stop codons. Following *in vitro* transcription and translation, the ribosome-displayed ORFeome can be screened *en masse* for binding to immobilized bait(s). Enrichment of candidate binding proteins can be rapidly assessed using quantitative real-time PCR (qPCR) with ORF-specific primers. Alternatively, proteomic-scale enrichment data can be obtained by performing deep sequencing of enriched libraries (Figure 1a). On current sequencing instruments, samples can be highly multiplexed, thereby bringing the cost of each proteomic screen well below $100. In addition, all steps required for PLATO can be fully automated with standard liquid handling robotics.

We considered multiple issues as we developed a strategy for deep sequencing of recovered display libraries. First, some amount of RNA transcript degradation occurs during ribosome display; we therefore avoided strategies dependent upon recovery of the full-length transcript, which would bias toward shorter ORFs. Exclusive recovery of the ORF 3′ termini ensures that the number of sequencing reads is stoichiometrically correlated with transcript abundance, which simplifies data analysis and minimizes the required sequencing depth. We therefore adopted the following protocol: after chemical fragmentation of enriched mRNAs and reverse transcription with a common primer, cDNA polyadenylation is followed by a two-stage PCR amplification to add sample barcodes and sequencing adapters (Figure 1b).

Subsequent multiplex deep sequencing analysis of pooled display libraries is reproducible and quantitative (Supplementary Figure 2). Sequencing an aliquot of unenriched human pRD-ORFeome mRNA ("input") detected the transcripts of 14,582 unique ORFs (Figure 1c). As examples of typical experimental applications, the ability of PLATO to identify known and novel protein interactions was assessed using 1) purified protein bait, 2) a complex mixture of patient autoantibodies, and 3) a small molecule kinase inhibitor.

Tyrosine-protein kinase LYN is a member of the Src family of non-receptor protein tyrosine kinases. This protein contains common structural components of Src family kinases, including N-terminal SH3 and SH2 domains, and a C-terminal kinase domain.[6] LYN has been extensively characterized for its interaction partners, making it an ideal bait to test the ability of PLATO to identify known interactions. After affinity enrichment of the human ORFeome using as bait GSTLYN, GST alone, or an irrelevant GST-fused protein (GST-Muted), we performed Illumina sequencing to identify proteins specifically precipitated with GST-LYN (Figure 2a, Supplementary Figure 3a). We noted that a number of established LYN binding partners were enriched specifically by GST-LYN, and validated two of these by qPCR (Figure 2b).[7, 8] We prioritized candidate ORFs for further investigation by their degree of enrichment on GST-LYN (Supplementary Table 1). GSTLYN precipitation and western blot analysis confirmed binding for five of seven novel candidates tested (Figure 2c). Of the two candidates that did not validate, one bound nonspecifically to GST, whereas the other was a true negative at the level of detection by western blot. Among the most highly enriched ORFs, protein domain analysis revealed that SH2 domain-containing proteins were overrepresented ($P = 0.0098$). Consistent with a role for LYN autophosphorylation in mediating these SH2 interactions, phosphatase pre-treatment of immobilized GST-LYN abolished binding of SH2D1A and SH2D4A, but only partly diminished binding of PIK3R3, suggesting the presence of an additional interaction domai (Supplementary Figure 3b). These proteins have not previously been reported to interact with LYN.

We next asked whether PLATO could be used to identify protein targets of antibodies from patients with autoimmune disease. As a preliminary control, we examined target enrichment due to affinity purified, commercially available antibodies. Anti-P53 and anti-PDCD4 antibodies were separately immobilized on protein A/G beads and used for immunoprecipitation of the displayed ORFeome library. By qPCR, P53 and PDCD4 were robustly enriched by their cognate antibodies, but not by the negative control IgG and RBM15 antibodies (Supplementary Figure 4).

In previous work, we synthesized an oligonucleotide library encoding a 36 residue overlapping human peptidome for display on bacteriophage T7 (T7-Pep). Deep sequencing of affinity enriched T7-Pep using autoimmune cerebrospinal fluid (CSF) from three individuals with paraneoplastic neurological disorder (PND) uncovered known and novel autoantigens.[9] We screened antibodies from these same three patient samples immobilized on protein A/G beads using the PLATO technology and compared the results. T7-Pep is a complete proteomic collection of 36 residue peptides, whereas the human ORFeome is a partial collection of full-length proteins, and our findings reflect this inherent complementarity. For example, neuro-oncological ventral antigen 1 (NOVA1) is absent

from the human ORFeome v5.1, and so PLATO was unable to detect this known autoreactivity in patient A, whereas it was robustly identified with T7-Pep. Conversely, PLATO identified numerous autoantigens for each patient that were missed in our peptidome screens (Supplementary Table 2). For example, PLATO analysis of patients A and B revealed immunoreactivity with known cancer autoantigens not detected with T7-Pep. Several of these reactive antigens were selected for confirmation via immunoprecipitation and western blotting (Figure 2e, Supplementary Figure 5b-d). In addition, we had previously established that antibodies from patient C recognized the tripartite motif containing proteins TRIM9 and TRIM67. PLATO considerably expanded the members of the TRIM family recognized by antibodies in this patient's CSF to include TRIM1/MID2, TRIM18/MID1, TRIM54, and TRIM55 (Figure 2d). Interestingly, multiple sequence alignment results in tight clustering of this precise subset of the extended TRIM family, suggesting the presence of shared, conformational epitopes there were not represented in T7-Pep.[10] As an alternative PLATO readout, microarray hybridization revealed a similar spectrum of enriched antigens (Supplementary Figure 6). The combination of multiple unbiased autoantigen discovery approaches, including PLATO, may eventually permit the saturating characterization of autoantibody repertoires.

Current small molecule target discovery typically involves the use of cell extracts and mass spectrometry, which suffers from the wide distribution of protein abundances and poor sensitivity. Normalized ORF libraries and quantitative DNA sequencing might therefore offer greater power to detect protein targets of small molecules. We tested this idea with gefitinib, a well-defined inhibitor of epidermal growth factor receptor's (EGFR) tyrosine kinase domain. Gefitinib interacts not only with the ATP-binding pocket of EGFR, but with other tyrosine kinase domains as well.[11] Candidate analysis after ORFeome affinity enrichment on gefitinib-coupled beads revealed significant enrichment of 10 out of the 17 predicted targets tested (Figure 2f). This experiment demonstrates the relative ease by which candidate protein interactions can be assayed with PLATO; the binding of any ORF can be rapidly assessed using qPCR and does not require cloning, protein purification, or western blotting. ORFeome libraries affinity enriched by the Src family tyrosine kinase inhibitor dasatinib exhibited overrepresentation of protein kinases ($P = 0.0003$), including the known target LCK and several not previously associated with this compound (Supplementary Table 3).

The experiments described here demonstrate the utility of the PLATO method, but it is also important to consider its limitations. As mentioned above, the human ORFeome collection is incomplete and such collections are not yet available for all organisms. However, the quality, completeness and availability of these libraries will continue to improve over time. As an *in vitro* method, in addition to lacking post translational modifications, large proteins may be expressed less efficiently, and proteins containing membrane-spanning or aggregation-prone domains that normally require host cellular machinery for proper folding may aggregate, complicating analysis. Ribosome display imposes certain limitations on the conditions under which affinity enrichments can be performed (e.g. low temperature and absence of RNAse contamination), and using proteins containing nucleic acid-binding domains as baits may result in very high non-specific binding, interfering with their use.

When the required conditions for PLATO are met, however, the method provides three main advantages as a tool for proteomic investigations. Compared to most alternative display platforms, ribosome display greatly diminishes protein size and composition constraints. Compared to protein microarray technology, the challenges and costs associated with protein purification, surface immobilization and microarray scanning are significantly circumvented. Finally, the rapidly declining cost of multiplex DNA sequencing will make PLATO an ideal platform for projects involving large numbers of screens, such as cohort-scale autoantibody profiling or structure-activity relationship analyses of small molecule compounds.

Methods

Plasmids and antibodies. A ribosome display (RD) backbone vector[5] was modified by inserting a Gateway cassette (attR1-ccdB-CM^R-attR2) to create the pRD-DEST destination vector according to the manufacturer's instructions (Invitrogen). Human ORFeome library v5.1 entry clones were pooled into eighteen super-pools (generally about ten 384-well entry plates per pool, based on plate serial number). For each super-pool, one LR reaction was performed to recombine the ORFs into the pRD-DEST vector. pDEST40 vector (Invitrogen) was used for transient expression of ORFs in 293T cells. pDEST15 vector (Invitrogen) was used for production of N-terminally GST-fused LYN (GST-LYN), MUTED (GST-MUTED) or control peptide (GST-Pep, DYKDDDDK) in BL21 E. coli. The antibodies used in this study include: rabbit IgG polyclonal antibody (Cat No. 2729, Cell Signaling), anti-p53 polyclonal antibody (Cat No. 9282, Cell Signaling), anti-PDCD4 polyclonal antibody (Cat No. A301-106A, Bethyl Lab), anti-RBM15 polyclonal antibody (Cat No. A300-821A, Bethyl Lab), and anti-V5 monoclonal antibody (Cat No. R960-25, Invitrogen).

Patient samples. PND cerebrospinal fluid samples were described previously.[9]

Ribosome display of human ORFeome v5.1. T7B and TolAk (0.2 μM) primers were used to PCR amply the pRD-ORFeome template (50 ng) by PrimeSTAR® HS PCR Kit (Takara). The following thermal cycles were used: 98 °C, 2 min/ 98 °C, 10 sec; 55 °C, 10 sec; 72 °C, 8 min; 10 cycles. The PCR product was purified using QIAquick® PCR Purification Kit (Qiagen). The purified PCR products of all eighteen super-pools were combined together in equal amounts (by weight). *In vitro* transcription was performed using T7 Ribomax Large *in vitro* Transcription Kit according to the manufacturer (Promega). RNA products were purified using RNeasy® column (Qiagen). *In vitro* translation was performed using RTS 100 E. coli HY Kit (5Prime). 7.5 μg mRNA in a 50 μl reaction containing 1 μl RNAseOUT (Invitrogen) was subjected to *in vitro* translation performed on PCR machine at 30 °C for 15 min. 12.5 μl aliquots of each translation reaction was diluted with 85.5 μl ice cold RD selection buffer (RD wash buffer (50 mM Tris Acetate, 150 mM NaCl, pH to 7.5 50 mM Mg Acetate, 0.5% Tween 20), 2.5 mg/ml heparin, 1% BSA, 100 μg/ml yeast tRNA with 2 μl RNAseOUT. The reaction mixture was centrifuged at 14,000 × g for 5 min at 4 °C. The supernatant was then moved to a new, ice cold tube.

ORFeome precipitation. GST-protein coated bead preparation: Expression of GST-Pep, GST-LYN and GST-MUTED was induced with 0.1 mM IPTG at 30 °C for 4 hours. Cells

were pelleted and lysed in lysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM EDTA, 5 mM $MgCl_2$, 0.2% NP-40, 2 mM DTT, 0.2 mM PMSF, 1 μg/ml pepstatin, 1 μg/ml aprotinin, 1 μg/ml leupeptin, 200 μg/ml lysozyme) on ice for 1 hour. The lysate was sonicated for 1 min on ice (Branson; output 4.0, duty cycle 50%). The lysate was centrifuged and supernatant retained. MagneGST™ glutathione particles (Promega) were coated with lysate at 4 °C for 4 hours. Beads were washed with buffer I (50 mM Tris pH 7.5, 500 mM NaCl, 1 mM EGTA, 10% Glycerol, 0.1% TritonX-100, 0.1% beta-mercaptoethanol, 1 mM PMSF) three times and buffer II (50 mM HEPES pH 7.5, 100 mM NaCl, 1 mM EGTA, 10% Glycerol, 0.1% beta-mercaptoethanol, 1 mM PMSF) three times. To assess LYN phosphotyrosine binding dependence, 20 μl glutathione particles containing approximately 2 μg protein were treated with 400 units of Lambda protein phosphatase (NEB) in 1x Protein MetalloPhosphatases buffer (50 mM HEPES, 100 mM NaCl, 2 mM DTT, 0.01 % Brij 35 and 1 mM $MnCl_2$) at 30 °C for 30 min with agitation. Patient antibody coated bead preparation: PND patient cerebrospinal fluid (CSF) containing 2.0 μg of immunoglobulin or 2.0 μg of rabbit IgG in 400 μl 1x PBST (3.2 mM Na_2HPO_4, 0.5 mM KH_2PO_4, 1.3 mM KCl, 135 mM NaCl, 0.05% Tween 20, pH 7.4) containing 1% acetylated bovine serum albumin (BSA) (Cat No. B2518, Sigma-Aldrich) were incubated with protein A/G magnetic beads (Invitrogen) at 4 °C overnight. Beads were washed with RD wash buffer five times. Drug coated bead preparation: gefitinib was immobilized on magnetic beads using a procedure previously described for covalently attaching the drug to sepharose 6B beads.[13] Biotinylated dasatinib (biotin-dasatinib, 500 μM) was immobilized on 50 μl of Dynabeads® MyOne™ Straptavidin T1 beads (Invitrogen) by incubation in 1x PBST containing 10% DMSO at 4 °C overnight. An equal amount of biotin (Sigma) was immobilized on beads as negative control. For all bead types: Beads were next blocked with RD selection buffer at 4 °C for 2 hours. 100 μl of the ice cold RD selection buffer containing the translated ORFeome library was then incubated with the beads at 4 °C for 6 hours while rotating. For competition experiments, free biotin-dasatinib (100 μM) was pre-incubated with the translated ORFeome library at 4 °C for 2 hours prior to the incubation of the library with biotin-dasatinib beads. Beads were then washed six times with 500 μl ice cold RD wash buffer. After the final wash, ribosomal complexes were disrupted by resuspension in 50 μl EB20 elution buffer (50 mM Tris Acetate pH 7.5, 150 mM NaCl, 20 mM EDTA) containing 1 μl RNAseOUT while rotating at 37 °C for 10 min. The eluted mRNA was then purified using an RNeasy® column (Qiagen).

RT-qPCR analysis of precipitated ORFs. Eluted mRNA samples were reverse transcribed using the primer PRDREV (0.1 μM) and SuperScript® III Reverse Transcriptase according to the manufacturer. The primer pair targeting the 3′ end of the corresponding ORF (0.1 μM each) was used to measure its mRNA level with Platinum® SYBR® Green qPCR SuperMix (Cat No. 11744, Invitrogen) on a 7500 Fast PCR-System (Applied Biosystems) The following thermal cycles were used for qPCR: 95 °C, 1 min/ 95 °C, 5 sec; 60 °C, 30 sec; 40 cycles.

mRNA sample preparation and Illumina sequencing. Recovered mRNA samples were fragmented using NEBNext® Magnesium RNA Fragmentation Module (NEB). The reaction was performed in a preheated thermal cycler for 3 min at 94 °C. Fragmented mRNA was cleaned up using Spin-50 mini-column (USA Scientific) and subjected to reverse

transcription (RT) using SuperScript® III Reverse Transcriptase (Invitrogen) and the TolA100RT primer (0.1 μM). After RT, the primer was removed by Exonuclease I (TaKaRa) digestion at 37 °C for 30 min. The mRNA template was then removed by incubation with RNase H (Invitrogen) at 37 °C for 30 min. cDNA was purified using QIAquick® PCR Purification Kit (Qiagen). Polyadenylation of cDNA 3′ end was performed using a TdT reaction kit (Invitrogen) according to the manufacturer. The TdT product was purified using QIAquick® PCR Purification Kit. The 1st PCR was performed using 0.25 μl Herculase II Fusion DNA Polymerase (Agilent), TdT product as the template, and 0.2 μM reverse primer Adaptor-$(dT)_{24}$ in 25 μl volume. A single thermal cycle was performed: 95 °C, 2 min/ 50 °C, 1 min/ 72 °C, 7 min. The forward primer, 0.2 μM P5-PRDREV was then added in an additional 25 μl PCR mixture. The following thermal cycles were then performed: 95 °C, 2 min/ 95 °C, 20 sec; 55 °C, 30 sec; 72 °C, 1 min; 30 cycles/ 72 °C, 5 min. The 1st PCR product was purified using QIAquick PCR Purification Kit. The 2nd PCR was composed of 0.5 μl Herculase II Fusion DNA Polymerase, 100 ng 1st PCR product as DNA template, and 0.2 μM primers (forward, P5-PRDREV; reverse, INDEX primer). The following thermal cycles were then performed: 95 °C, 2 min/ 95 °C, 20 sec; 55 °C, 30 sec; 72 °C, 1 min; 10 cycles/ 72 °C, 5 min. The 2nd PCR product was purified using QIAquick PCR Purification Kit. The purified 2nd PCR product was quantified on a 7500 Fast PCR-System (Applied Biosystems) using Platinum® SYBR® Green qPCR SuperMix (Invitrogen), and 2 μl of 5 μM P5 and P7_2 mix. The following thermal cycles were used: 50 °C, 2 min/ 95 °C, 10 min/ 95 °C, 15 sec; 60 °C, 2 min; 35 cycles. An equal amount of each 2nd PCR product was combined and sequenced on the Illumina HiSeq 2000 using a 50 cycle, multiplex single-end protocol with the custom primer, PRDREV-attB2-SP.

Analysis of Illumina data. Sequences were aligned using the Bowtie software, version 0.12.7. An index file was constructed using the 50 3′-most nucleotides of each sequence from the ORFeome v5.1. Alignments were performed using the following parameters: -n 2 -l 30 --best --nomaqround --norc -k 1. A single mismatch was allowed in the 7 nucleotide barcode sequence which was used to assign each read to the appropriate sample library. We typically obtained between 5 and 10 million aligned reads per barcoded library. The alignments corresponding to each ORF were then aggregated, thus defining each library's read count vector. We considered only ORFs with an IP count greater than a certain threshold. Enrichment was then calculated by adding a pseudocount of 1 to each clone and then dividing the fractional abundance of each IP'ed clone by that in the appropriate negative control. For LYN IP's the negative control was GST-Pep or GST-MUTED, and for PND IP's the negative control was rabbit IgG. For biotin-dasatinib, the negative control was biotin.

Validation of candidate interactions. pDEST40 plasmid harboring candidate ORFs was transiently transfected into 293T cells using Fugene® HD transfection reagent (Promega) according to the manufacturer. 48 hours post transfection, cells were harvested in 1 ml of 1x RIPA buffer (50 mM Tris-HCl pH 7.4, 1% NP-40, 0.25% Na deoxycholate, 150 mM NaCl, and 1 mM EDTA) containing protease inhibitor cocktail (1 mM PMSF, 1 μg/ml aprotinin, 1 μg/ml leupeptin, and 1 μg/ml pepstatin) and phosphatase inhibitor cocktail (Sigma). Precipitation was performed by incubating either bait-coated glutathione or protein A/G beads with the 293T cell lysates at 4 °C overnight with rotating. Beads were washed with ice

cold 1x RIPA buffer six times and eluted in 2x Laemmli sample buffer. Samples were then separated on a SDS-PAGE gel. An anti-V5 antibody was used to detect the presence of candidate proteins after transfer onto a PVDF membrane.

Microarray hybridization. Eluted mRNAs IP′ed from mixed PND sample (patient A:patient C = 1:1) was reverse transcribed using the primer PRDREV (as above). PCR was performed using cDNA as template, attB1 and attB2 (0.2 μM) primers by PrimeSTAR® HS PCR Kit. PCR products were recombined into entry vector pDONR223 by BP reaction. After transformation into DH5α E. coli, entry clones were recovered and then recombined into the destination vector pRD-DEST with the LR reaction. After transformation into DH5α E. coli, LR clones were recovered and subjected to PCR using T7B and TolAk primers. PCR products were purified for T7 *in vitro* transcription using MEGAscript® T7 Kit according to the manufacturer's instructions (Ambion). *In vitro* transcribed PND affinity purified sample mRNAs were labeled with Cy3 dye. The total input *in vitro* transcribed ORFeome mRNAs were labeled with Cy5 dye. Cy3 and Cy5-labeled RNAs were mixed (1:1) and hybridized on our custom human ORFeome microarrays (Agilent).[14]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank K. Waraska, M. Cicero, S. Alian and A. Gagne for assistance with Illumina sequencing, and J. Laserson for statistical advice. Thanks to Dianrong Zhu for help with synthesis of biotin-dasatinib, which was partially supported by NIH U54 CA156734 to the University of Massachusetts Boston – Dana-Farber Harvard Cancer Center U54 Comprehensive Partnership (Project 3, Co-PIs: N. S. Gray, W. Zhang, and P. L. Yang). We also thank Nathanael Gray atHarvard Medical School for valuable advice, and Harold Varmus at the NCI for providing gefitinib reagents and advice. This work was supported in part by NIH grant 3P30CA023100-25S8 to S.K. S.J.E. is an investigator with the Howard Hughes Medical Institute.

References

1. Faix PH, et al. Phage display of cDNA libraries: enrichment of cDNA expression using open reading frame selection. BioTechniques. 2004; 36:1018–1022. 1024, 1026–1019. [PubMed: 15211753]
2. Zhu H, Bilgin M, Snyder M. Proteomics. Annual review of biochemistry. 2003; 72:783–812.
3. Jeong JS, et al. Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. Mol Cell Proteomics. 2012; 11:O111, 016253. [PubMed: 22307071]
4. Lamesch P, et al. hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. Genomics. 2007; 89:307–315. [PubMed: 17207965]
5. Amstutz P, Binz Zahnd, C. Pluckthun A. Ribosome Display: In Vitro Selection of Protein-Protein Interactions. Cell Biology, A Laborator Manual. 2006; 1:13.
6. Boggon TJ, Eck MJ. Structure and regulation of Src family kinases. Oncogene. 2004; 23:7918–7927. [PubMed: 15489910]
7. Weng Z, et al. Identification of Src, Fyn, and Lyn SH3-binding proteins: implications for a function of SH3 domains. Mol Cell Biol. 1994; 14:4509–4521. [PubMed: 7516469]
8. Samuels AL, Klinken SP, Ingley E. Liar, a novel Lyn-binding nuclear/cytoplasmic shuttling protein that influences erythropoietin-induced differentiation. Blood. 2009; 113:3845–3856. [PubMed: 19064729]
9. Larman HB, et al. Autoantigen discovery with a synthetic human peptidome. Nat Biotechnol. 2011; 29:535–541. [PubMed: 21602805]

10. Carthagena L, et al. Human TRIM gene expression in response to interferons. PLoS ONE. 2009; 4:e4894. [PubMed: 19290053]
11. Brehmer D, et al. Cellular targets of gefitinib. Cancer Res. 2005; 65:379–382. [PubMed: 15695376]
12. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009; 37:1–13. [PubMed: 19033363]
13. Somwar R, et al. Superoxide dismutase 1 (SOD1) is a target for a small molecule identified in a screen for inhibitors of the growth of lung adenocarcinoma cell lines. Proc Natl Acad Sci U S A. 2011; 108:16375–16380. [PubMed: 21930909]
14. Yen HC, Xu Q, Chou DM, Zhao Z, Elledge SJ. Global protein stability profiling in mammalian cells. Science. 2008; 322:918–923. [PubMed: 18988847]

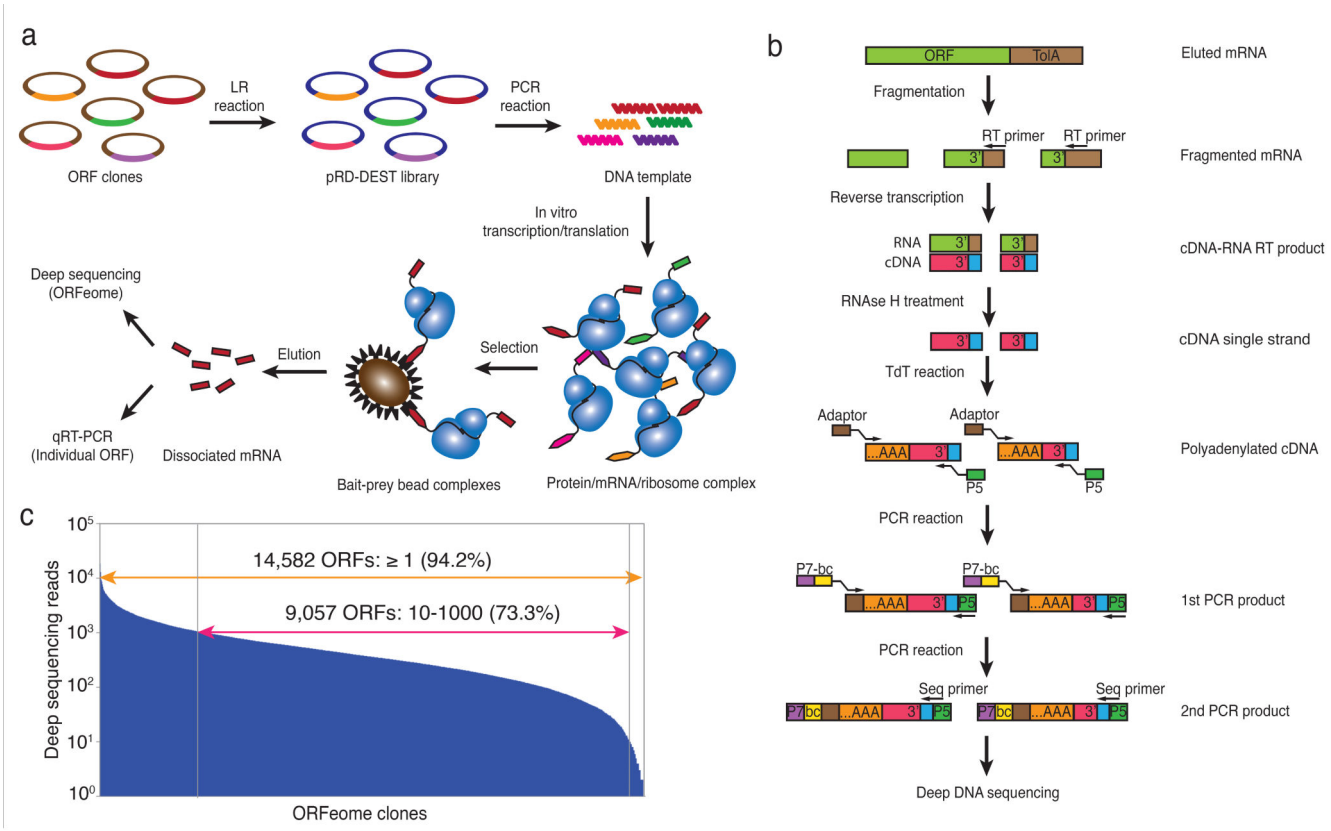

Figure 1.

Parallel Analysis of *in vitro* Translated ORFs (PLATO). (a) ORF display scheme. The pooled human ORFeome v5.1 entry vector library is recombined into the pRD-DEST expression vector. Expression plasmids are PCR amplified to generate the DNA templates for *in vitro* transcription. Following *in vitro* translation, the protein/mRNA/ribosome complexes are incubated with protein, antibody or small molecule bait immobilized on beads. The enriched mRNA library is finally recovered from bait-prey bead complexes for further analysis. (b) Processing of mRNA samples for deep DNA sequencing. After fragmentation and reverse transcription (RT) using a universal primer to recover the 3′ end of ORFeome transcripts, cDNA is polyadenylated with terminal deoxynucleotide transferase (TdT) and amplified for multiplex deep sequencing using primers containing a sample barcode, and the P5 and P7 Illumina sequencing adaptors. (c) Sequencing reads of the unenriched human pRD-ORFeome mRNA library (the “input” library). Most ORFs were sequenced at least once.

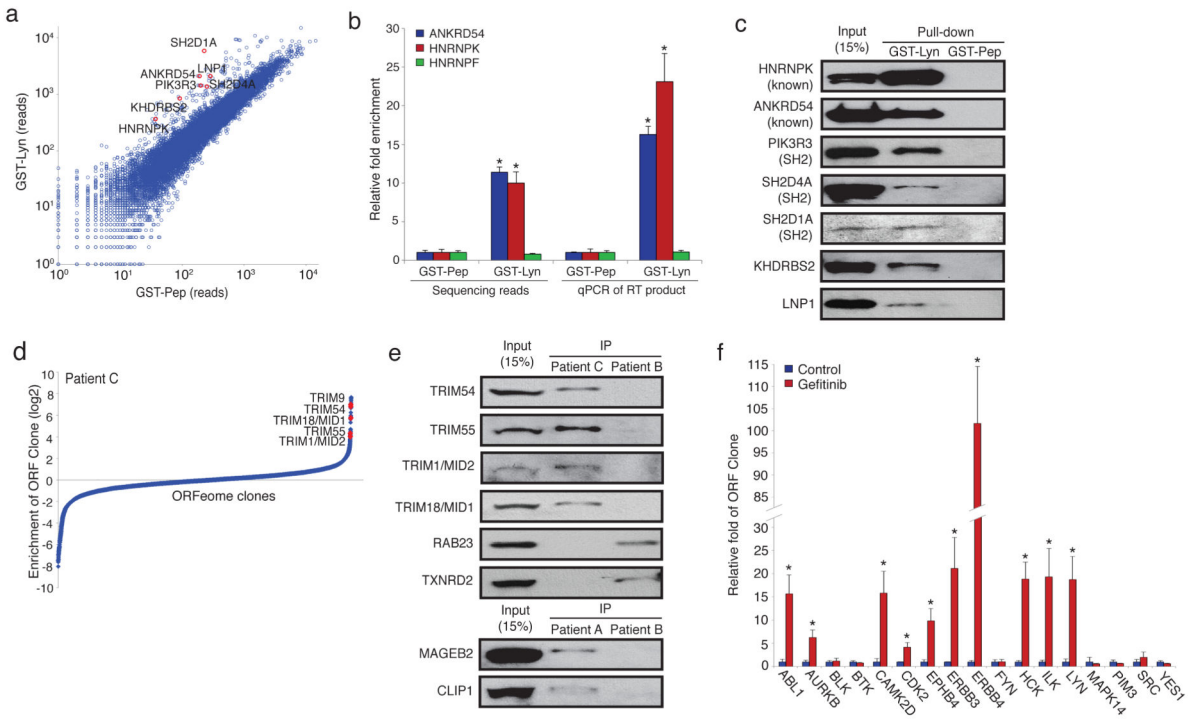

Figure 2.

Identification of known and novel interactions using PLATO. (a) Scatter plot of each ORF's sequencing reads after enrichment on GST-LYN or GST-Pep. Several known and novel LYN binding candidates are highlighted in red. (b) Enrichment of two known LYN interactors. Data was normalized to the GST-Pep enriched libraries (n=3, mean ± s.d.). * $P <$ 0.01 (c) Confirmation of known and predicted LYN binding proteins by affinity precipitation-western blotting of lysates from HEK 293T cells transiently overexpressing the individual V5-His-tagged candidate proteins. (d) Enrichment ranking of PND autoantigens identified using CSF from patient C. (e) Confirmation of previously unidentified PND patient autoantigens. (f) Enrichment of previously identified gefitinib targets. Data was normalized to the control-enriched libraries (n=3, mean ± s.d.). * $P < 0.05$.